BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 9: e178694
doi: 10.3897/biss.9.178694

OPEN ACCESS

Conference Abstract

Sounds of Life: A FAIR Platform to Unlock the Potential of ioacoustic Data for Discovery and Interoperability

Julien Homo‡, Kévin Darty‡, Shadia Kilouchi§

‡ Foxcub, Issy-Les-Moulineaux, France
§ CNRS UMR Passages, Pessac, France

Corresponding author: Julien Homo (julien.homo@foxcub.fr), Kévin Darty (kevin.darty@foxcub.fr), Shadia Kilouchi (shadia.kilouchi@cnrs.fr)

Received: 19 Nov 2025 | Published: 24 Nov 2025

Citation: Homo J, Darty K, Kilouchi S (2025) Sounds of Life: A FAIR Platform to Unlock the Potential of ioacoustic Data for Discovery and Interoperability. Biodiversity Information Science and Standards 9: e178694. https://doi.org/10.3897/biss.9.178694

Abstract

Context and motivation

Ecoacoustics is an emerging discipline that studies environmental sounds across spatial and temporal scales to understand biodiversity and ecosystem dynamics, from species behaviour to landscape-level change (Sueur and Farina 2015).

Soundscapes encompass biological, abiotic, and anthropogenic acoustic sources, providing a comprehensive and dynamic view of ecosystems. Despite the increasing abundance of recordings collected worldwide, environmental acoustic data remain scattered, inconsistently described, and difficult to reuse due to the absence of shared international standards and interoperable workflows (Sugai et al. 2019).

To unlock this latent wealth into actionable biodiversity knowledge, structuring and interoperability are essential.

© Homo J et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

The Sounds of Life initiative

Launched in 2023, Sounds of Life (SoL) unites over a dozen French institutions bridging ecology, ecoacoustics, geography, and information sciences. Coordinated by CNRS-UMR Passages and labelised by Huma-Num, the French national infrastructure for digital humanities, SoL aims to build a shared FAIR framework (Findable, Accessible, Interoperable, Reusable) for environmental acoustic data. The initiative brings together research groups across France such as CEFE, ISYEB, CESCO, INRAE, Lab-STICC, and others, combining biodiversity expertise with data science and infrastructure design.

At the heart of the project is a FAIR-by-design platform that operationalizes the full pipeline: ingest, enrich, annotate, and publish. Through a web-based user interface (Homo et al. 2025), users can import audio files from local or cloud sources. A CSV mapping step lets users attach a first layer of structured metadata on upload, while built-in validators check required and recommended fields for GBIF-ready publication. These imported records are shared in a library, enabling collaboration in workspaces. Vocabulary services power autocomplete and normalization. A customizable spectrogram viewer Fig. 1 enables users to annotate precise time-frequency regions, add comments and provenance. Sets of records, called datasets, can be defined with mappings to Darwin Core terms and packaged as Darwin Core Archives (DwCA) with GBIF compliance indicators. Support for persistent identifiers in publication workflows is planned for the near future.

Lessons learned from the pilot

Building SoL around the FAIR principles revealed key lessons in representing acoustic data within biodiversity standards such as Darwin Core. A first challenge concerned multi-species soundscapes, which are difficult to represent within models that assume a single occurrence per recording. While Darwin Core handles single-taxon recordings well, real soundscapes often include several species and non-biological sounds. In our GBIF/TDWG-aligned profile, each recording is modeled as an Event and each detection as an Occurrence linked to that Event, with the media held via the Simple Multimedia extension. Time-frequency annotations are stored alongside detections, and exports reference the same media record. This design remains GBIF-compatible but reveals friction when publishing dense, multi-species scenes, highlighting the need for a community pattern or lightweight extension supporting multi-label audio.

Another key finding concerned vocabulary load. Beyond the about 190 Darwin Core terms, more than 60 additional fields from 10 domain vocabularies were required for taxonomy, geography, devices, and traits. To keep forms usable, we applied progressive field display strategy (required or optional), use-case presets, and GBIF-ready validation. A minimal acoustic profile (about 25 fields) complemented by an extended layer could provide a practical balance between expressiveness and usability.

We also faced structural challenges with the lack of a central schema enforcing consistency between Event, Occurrence, and MeasurementOrFact cores. We added

consistency rules and export checks to reduce hidden errors. Documenting such patterns in a concise TDWG guidance note would help others align acoustic workflows with existing biodiversity standards.

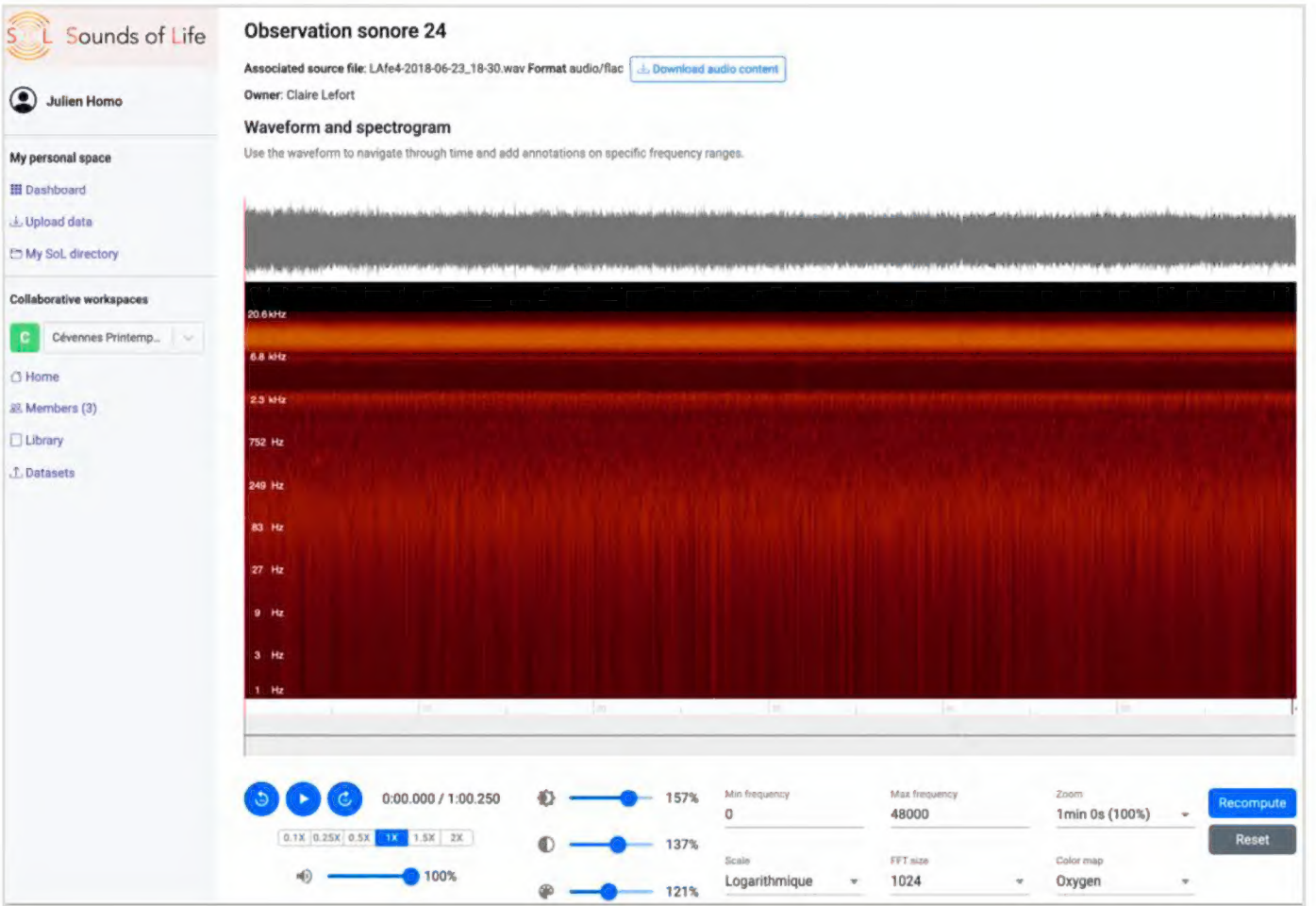

Figure 1.

SoL web interface for acoustic data exploration. Screenshot from the Sounds of Life platform (Julien Homo & Sounds of Life consortium, 2025), licensed under CC BY 4.0.

Overall, these lessons underline that interoperability is as much about design and usability as it is about technical standards. Shared profiles, cross-field validation rules, and clear community guidance will make FAIR acoustics both achievable and sustainable at scale.

Collaboration, AI-readiness, and next steps

Beyond technical design, SoL enables collaborative curation between ecologists, data specialists, and sound engineers, aligning vocabularies and improving reproducibility across studies, supporting the growing use of environmental sound as a measurable ecological signal.

While the current pilot focuses on metadata management and annotation, it is intentionally AI-ready. Consistent, well-structured, and traceable data are prerequisites for automated detection, soundscape classification, and large-scale monitoring. Recent syntheses (e.g., Hoefer et al. (2023)) in passive acoustic monitoring emphasize the need for standardized, comparable datasets to enable robust cross-site analyses.

Looking ahead, SoL aims to scale carefully beyond its pilot phase and contribute back to the TDWG/GBIF community through shared profiles and reference datasets. In the short term, through mid-2026, we are seeking input from TDWG and GBIF experts, ecoacoustics researchers, and data publishers to

1. review our Darwin Core profile for acoustic data (Event/Occurrence + Multimedia) and controlled vocabularies for sound types and scenes;
2. test end-to-end exports (DwC-A to IPT/GBIF) using real multi-species soundscapes;
3. contribute pilot datasets across biomes to evaluate FAIRness, traceability, and reusability; and
4. co-design AI-ready benchmark corpora with clear licensing, balanced labels, and robust metadata.

Our goal is to align these efforts with European and global infrastructures so that acoustic evidence becomes durable and comparable at scale.

Keywords

ecoacoustics, soundscape ecology, biodiversity monitoring, environmental recordings, metadata interoperability, data standardization, controlled vocabularies, GBIF integration, Darwin Core, acoustic indices, collaborative audio curation, spectrogram annotation

Presenting author

Julien Homo

Presented at

Living Data 2025

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Hoefer S, McKnight D, Allen-Ankins S, Nordberg E, Schwarzkopf L (2023) Passive acoustic monitoring in terrestrial vertebrates: a review. Bioacoustics 32 (5): 506-531. URL: https://doi.org/10.1080/09524622.2023.2209052
- Homo J, Darty K, Kilouchi S (2025) Sounds of Life: a FAIR platform to unlock the potential of acoustic biodiversity data . Living Data 2025. Video.

- Sueur J, Farina A (2015) Ecoacoustics: the Ecological Investigation and Interpretation of Environmental Sound. Biosemiotics 8 (3): 493-502. https://doi.org/10.1007/s12304-015-9248-x
- Sugai LSM, Desjonquères C, Silva TSF, Llusia D (2019) A roadmap for survey designs in terrestrial acoustic monitoring. Remote Sensing in Ecology and Conservation 6 (3): 220-235. https://doi.org/10.1002/rse2.131